Gujarati: Script of Western India

Designing and building a Gujarati sans-serif typeface family

2013-2015, Prof. Gary Munch, Senior Lecturer in Design, SASD at UB

ગુજરાતી લિપિ

g-u ja r-a tii l-i p-i

Gujarat State in West Central India (salmon pink) [Wikipedia]

GUJARATI IS A BRAHMIC SCRIPT of the western Indian subcontinental area, Gujarat. About 30 million people use it daily for the language Gujarati, and for Sanskrit for religious and scholarly purposes. It can also be used to write the Parsi-based language of the now-dwindling population of Zoroastrians who immigrated from Persia after the Islamic influx, and now read and write Gujarati as their daily script.

Gujarati is a fairly recent (four or five centuries) development from its immediate parent script, Devanagari. Devanagari is the standard script used for historical Sanskrit and for Hindi, the majority language of India.

Development of Gujarati from Devanagari

म→મ

Gujarati merchants and bankers developed a set of quicker, simpler forms for the Devanagari letters. In comparison the underlying forms of the letters are almost identical in most of the characters' structures. The one obvious difference is the removal of Devanagari's characteristic head stroke; so, Dev. अ *a* becomes Guj. અ *a* ; D. न *na* becomes G. ન *na*; D. म *ma* becomes G. મ *ma*. But as if in a conservation of detail, it is traditional (though not absolutely necessary) to include an exit serif at the lower right end of the letter stem. The exit stroke is a mark of Gujarati's origin as a fast, semi-cursive hand.

Abugida: a syllabary with a single inherant vowel

As with all Brahmic scripts, the Gujarati (and Devanagari) script is an abugida, a type of syllabary. Each basic letter represents an inherent vowel attached to the consonant. About 53 basic characters represent the Gujarati consonants plus *a* (actually [ə], schwa). So, મ *ma* can stand for not just the nasal sonorant *m* but also for the syllable *ma*.

Representing a variety of vowels

ા ો ો ૌ

To build a representation of a syllable like [*mu*], a separate vowel sign, or diacritic, is added to the base letter. So, મુ [*mu*], is built from મ *m* + ુ *u* diacritic. Vowel diacritics replace the value of the inherent schwa.

Freestanding vowel signs are available for all vowels that can begin a word. But within a word, vowels are indicated instead with diacritics placed over, under, or next to the basic consonant + schwa. The diacritic replaces the schwa with the diacritic's vowel-value. That's all very simple to accomplish with minimal positioning techniques that use the text stream sequence for ordering the letter glyphs. But, there's a catch: one of the vowel signs is placed not above, below, or after the consonant sign but before it.

Short i (ih) prepositioned to consonants

િ

The Gujarati િ, or simple [*i*] (English *ih*), must visually precede a consonant or consonant cluster; મ *ma* becomes [*mi*] not as મિ *m+i,* but as િમ *i+m*, and is then read as *mi*. The intelligence to do this trans-positioning must be built into the typeface font software and into the application software (and often the operating software). In addition, the િ is expected to adjust the length of its arc to accommodate the width of a ligature or conjunct cluster. That adjustment means multiple forms — regular, long, longer, etc. — are needed for correct typesetting.

Conjuncts: consonants cluster constantly

ન→ન્

Two ligature (letters tied, or ligated, together) glyphs in Gujarati are mandatory to express two common consonant clusters: ક્ષ *kssa* and જ્ઞ *jña*.

To build multiple-consonant clusters such as *kl* or *ndr* or *nstr,* Gujarati (and the other Devanagari-descended scripts) use a complex system of partial forms that combine into collections called conjuncts. Since the standard forms of most Gujarati letters have a stem, the half-forms of most letters are, like the ્, additional to the basic set of glyphs. The half-forms reveal the handwritten nature of the script; it is a simple matter to omit a stem in writing, and to connect the significant part of the first glyph to the following glyph. So, a cluster of sounds like અનમ *a+na+ma*, when intended to be read as *anma*, would be constructed from the horizontal of *na* ન and the full form of *ma* મ as અન્મ. Half-forms nearly double the glyph complement for Gujarati.

R and N reduction

મ્રમ્ર

When the *r* ર is not the starting part of a conjunt, the *r* ર takes a special reduced form as a small diagonal (called a *rakaar*) and attaches to the stems of most glyphs; *mra* મર becomes મ્ર, and *sra* સર becomes સ્ર, and *stra* સ્તર becomes the neat and compact સ્ત્ર (not to be confused with vocalic *r*, ઋ). Ocassionally the *rakaar* is added at the left: *drya* દ્રય is one glyph, દ્ય, where the *ya* ય glyph is reduced to a stem and flourish on *da* દ, with the *rakaar* — the middle sound of the three-consonant cluster — set at the left of its companions! As well, *r* can, as the first component of a conjunct, be a small curl ર્ (called *reph*) at the far right of the base form; *rshca* ર્શ્ચ.

Complement of Conjuncts

છ્ત્ર

Most fonts use a pre-composed form for these essential conjuncts with *r*, and a similar treatment of the *n* ન; *agn* અગ્ન could be written અગ્ર, or *issn* ઈષ્ન as ઈષ્ણ. Other nasals can be written with *anusvara* ં or *candrabindu* ઁ — as in *aum* ૐ.

With forms that have no stem, the *r* is added as a small angle underneath; so, *ttra* ટ્ર (*tta* ટ having no stem) becomes ટ્ર, or *chra* છ્ર becomes છ્ર; *ssttra* ષ્ટ્ર becomes a single form, ષ્ટ્ર. Or, the *r* stroke is tucked up inside of the base form; *kra* ક્ર becomes ક્ર, *tra* ત્ર is ત્ર, *jra* જ્ર is જ્ર. All of these need separately drawn forms.

Reversion to earlier forms

શ્ચ

In other conjuncts the base consonant might take a form older than Gujarati, and revert to a more Devanagari form. In some clusters with *sha* શ, the શ reverts to a tail–less curl, �improvised, from Devanagari *sha* श that has more forms added to it: *shca* શ્ચ (*ca* ચ reverts to D. च), *shra* શ્ર, *shna* શ્ન, *shla* શ્લ, *shva* શ્વ. Similarly, in conjuncts the *ha* હ can revert to its headless core Devanagari form ह, ह: *hnna* હ્ન, *hna* હ્ન, *hla* હ્લ, *hva* હ્વ, *hma* હ્મ, and *hya* હ્ય.

Stacking forms

ઢન→ઢ્ન

In some conjuncts the forms stack one over the other. So, *jja* જ્જ is જ્જ, *ngka* ઙ્ક is ઙ્ક, *ngga* is ઙ્ગ, *ddva,* ડ્વ; *ddhddha,* ઢ્ઢ; *ddhna,* ઢ્ન. And of course some stacked forms use Devanagari forms, as well: *nyca* ઞ્ચ is composed from *nya* ઞ and *ca* ચ (a soft *tsha* sound), with the D. form of *ca,*च, and ञ is *nyja,* from *nya* ઞ above and *ja* જ below, in its D. form of *ja* ज. Some can dangle at an angle off of *da* દ: *da gha* દ્ઘ, *da bha* દ્ભ, etc.

Kill the vowel!

શ્ચ્

If a vowel sound needs to be removed from a sound cluster and there is no initial half-form or conjunct available, a stroke called *virama* or *halant* is used: *cha ta* would be written *cha* + *virama* + *ta* છ્ત. But in many situations, the loss of the vowel(s) in a cluster is understood and is accounted for in reading by experience.

30 million and going strong

ધન્યવાદ!

Gujarati and its cousin scripts are formidable challenges to those who are accustomed to more linearly-organized scripts (Latin, Greek, Hebrew, Arabic (which latter has its own mandatory and optional complexities), Cyrillic, or Armenian. But, for all its complexity, Gujarati is written, composed, and read by millions who, after learning, have no more difficulty than users of other scripts of the world.

માનવ અધિકારોની વિશ્વવ્યાપી ઘોષણા

પ્રતિષ્ઠા અને અધિકારોની દૃષ્ટિએ સર્વ માનવો જન્મથી સ્વતંત્ર અને સમાન હોય છે. તેમનામાં વિચારશક્તિ અને અંતઃકરણ હોય છે અને તેમણે પરસ્પર બંધુત્વની ભાવનાથીનવર્તવું જોઈએ.

First Article of the UN *Universal Declaration of Human Rights* — All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.

AT TYPECON 2013 IN PORTLAND, OREGON, a workshop was offered by Kalapi Gajjar-Bordawekar, a designer at Dalton Maag in London. The topic was his native script, Gujarati. Having heard and seen the complexity of Devanagari script, it seemed possible that, as a simplified version of Devanagari, there wouldn't be so much need in Gujarati for the complicated manipulations that Devanagari requires.

All went well through the morning, as we stepped through the vowels and consonants. We worked with edged pens to experience the way the stress pattern — the placement of thick and thin in a letterform — differs from that of the European script. In Gujarati the axis is from upper right to lower left; in European scripts it is usually from upper left to lower right. After years of ensuring this major aspect in European letterforms, it was hand-and-mind boggling to switch! OAS to વઽટ! We talked about and adjusted needed proportions, boldness and lightness, angles and joins: all the meat of typeface design. And the curious behavior of the pre-positioned letter i. That was a tip off, but ... off to lunch!

Then the Boom. So, says Kalapi, that's that and by the way, 53 basic shapes, so now let's look at conjuncts (OMG there it was, the elephant had been there all the time, we were hooked, too late, no escape, down the rabbit hole of Conjuncts!). Upwards of 5000 possible, but with typically fewer for most designs, conjuncts come in a variety of forms and usages. Half forms to join to full forms, half forms pre-joined to other half forms, special positional versions, inset strokes and reverting to ancestral forms, stacked or .... Now, two years later, the basics for a Gujarati set of typefaces that only just need the smarts of OpenType substitution and positioning to function in a keyboarding environment... and multiple versions with different lengths of the *i* િ matra forms, which must reach to the rightmost edge of the glyph or conjunct they arch over... and combinations of diacritic vowels with each other and reph, rakaar, anusvara...still in progress. Fun! ધન્યવાદ, Kalapi!

ક્ષ ક્ક્ષ્ય જ્ઞ જ્ઞ જા જી જા જી i ણુ રૂ રુ દૃ દ્ર ઋ ક્ર ખ્ર ગ્ર દ્ર ઞ્ચ જ્ર ઞ્ર ણ્ર ત્ર થ્ર દ્ર ધ્ર ન્ર પ્ર ફ્ર બ્ર ભ્ર મ્ર ય્ર લ્ર વ્ર શ્ર સ્ર હ્ર ક્ક ક્ટ ક્ત ક્મ ક્ય ક્શ ક્સ ક્લ ક્વ ક્ષ ગ્ગ ઘ્ઘ ઙ્ક ક્ષ્ય ક્ષ્ય ક્ષ્ય